Bronchoalveolar lavage fluid lymphocytosis in chronic hypersensitivity pneumonitis: a systematic review and meta-analysis

Nicola Adderley, Christopher J. Humphreys, Hayley Barnes, Zahra A. Premji, Brett Ley, and Kerri A. Johannson

Affiliations: 1Faculty of Medicine, University of Calgary, Calgary, AB, Canada. 2Dept of Medicine, University of Calgary, Calgary, AB, Canada. 3Dept of Allergy, Immunology and Respiratory Medicine, Alfred Hospital, Melbourne, Australia. 4Dept of Medicine, University of California, San Francisco, San Francisco, CA, USA. 5Depts of Libraries and Cultural Resources, University of Calgary, Calgary, AB, Canada. 6Community Health Sciences, University of Calgary, Calgary, AB, Canada.

Correspondence: Kerri A. Johannson, 4448 Front Street S.E., Calgary, AB, T3M-1M4, Canada. E-mail: kerri.johannson@ahs.ca


ABSTRACT

Background: The role of bronchoalveolar lavage fluid (BALF) lymphocyte percentage in diagnosing chronic hypersensitivity pneumonitis (CHP) is unclear. We conducted a systematic review and meta-analysis of bronchoalveolar lavage (BAL) lymphocyte percentage in the diagnosis of CHP.

Methods: We searched Medline, Embase and the Cochrane Library from inception to August 2019. Individual patient data were obtained to test performance characteristics of BAL lymphocyte percentage at different thresholds. Random-effects models were used for pooled estimates, with comparisons made between CHP and non-CHP interstitial lung diseases (ILDs).

Results: Fifty-three studies were included in the systematic review and 42 in the meta-analysis. The pooled estimate for BAL lymphocyte percentage was 42.8% (95% CI 37.7–47.8, I²=95.3%) in CHP, 10.0% (95% CI 6.9–13.1, I²=91.2%) in idiopathic pulmonary fibrosis (IPF), 23.1% (95% CI 3.0–43.2, I²=85.2%) in non-IPF idiopathic interstitial pneumonia (IIP), 23.4% (95% CI 11.0–35.9, I²=45.7%) in connective-tissue disease associated ILD (CTD-ILD) and 31.2% (95% CI 17.6–44.8, I²=95.2%) in sarcoidosis. Results differed between CHP and IPF (p<0.0001), non-IPF IIP (p=0.0309) or CTD-ILD (p=0.0824), but not between CHP and sarcoidosis (p=0.0966). Using individual patient data from eight studies, a lymphocyte percentage threshold of >20% provided a sensitivity of 68.1% and a specificity of 64.8% for CHP. Higher thresholds provided lower sensitivity with higher specificity. Older age and ever having smoked were associated with lower lymphocyte percentage in CHP.

Conclusions: BAL lymphocyte percentage is higher in CHP compared to IPF and other IIPs, with higher thresholds providing improved specificity at the cost of sensitivity. However, the parent studies are at risk of incorporation bias and prospective studies should evaluate the additive discriminate value of BAL lymphocyte percentage to accurately diagnose CHP.
Introduction

Hypersensitivity pneumonitis (HP) is an inflammatory and/or fibrotic immune-mediated interstitial lung disease (ILD) caused by sensitisation to an inciting antigen. In its chronic form, HP is characterised by an insidious progressive course that obscures the link between causative antigen and disease. Chronic hypersensitivity pneumonitis (CHP) shares overlapping clinical and radiological features with other ILDs [1] and the absence of consensus diagnostic criteria further complicates establishing a diagnosis. Differentiating CHP from non-CHP ILDs (e.g. idiopathic pulmonary fibrosis (IPF) and idiopathic nonspecific interstitial pneumonia (NSIP)) can be challenging, but is critical for disease management and prognostication [2, 3]. The gold standard for diagnosis of CHP involves multidisciplinary team discussion and integration of radiological, clinical and pathological data [4].

Bronchoalveolar lavage fluid (BALF) analysis is proposed as an informative tool in the diagnostic evaluation of patients with HP and CHP [5]. Increased cellularity with lymphocytosis is associated with HP and the range of lymphocyte counts is believed to reflect the degree of alveolitis [6, 7]. However, there is a paucity of robust evidence supporting the role of alveolar lymphocytosis in diagnosing CHP [8]. In a study of antigen-determinate HP patients, the mean lymphocyte percentage was elevated in all forms of disease, but was lower in CHP compared to the acute or subacute forms [9]. Radiographic fibrosis is associated with lower bronchoalveolar lavage (BAL) lymphocyte percentage in ILD [10] and, in advanced fibrotic HP with a histological pattern resembling usual interstitial pneumonia (UIP), BAL lymphocytosis may be less pronounced [11]. In a recent Delphi study, more than 75% of ILD experts rated BAL lymphocytosis >40% as “important” or “very important” for the diagnosis of CHP [4]. No consensus was reached on the importance of BAL lymphocytosis in the 30–39% range and findings of 20–30% were deemed uninformative. These findings underscore the need for research to identify an optimal threshold for BAL lymphocytosis in the diagnosis of CHP.

BAL lymphocytosis may be influenced by several variables, including the presence and extent of fibrosis, the timing relative to antigen exposure, smoking status and the procedural technique for BAL collection [5]. In addition, the presence of BAL lymphocytosis may not differentiate between other histologic entities also characterised by lymphocytic inflammation, particularly NSIP or cryptogenic organising pneumonia (OP) [12]. International guidelines exist to guide the BALF collection procedure, yet there remains heterogeneity in collection, processing and analysis and there remains the potential for misclassification [13, 14]. BAL lymphocyte subset analysis (i.e. CD4:CD8 ratio) was historically thought to be helpful in establishing diagnoses of specific ILDs; however, recent data suggest it is not informative in CHP and testing is not routinely recommended in ILD [5].

The role of BAL lymphocytosis in establishing a diagnosis of CHP remains unclear. The aim of this systematic review and meta-analysis was to describe BAL lymphocyte percentage in CHP and compare these findings to non-CHP ILDs. We further sought to test the performance characteristics of BAL lymphocyte percentage at different thresholds to accurately differentiate CHP from other non-CHP ILDs.

Methods

Search strategy and selection criteria

We performed a systematic review and meta-analysis following Meta-analysis Of Observational Studies in Epidemiology (MOOSE) guidelines [15]. The protocol was registered in the PROSPERO database (CRD42019122236). Searches of Medline, the Cochrane Central Register of Clinical Trials and Embase were conducted from database inception to August 2019. The search strategy design was supported by an academic research librarian (Z.A. Premji) and included both text words and controlled vocabulary (see supplementary table S1 for details). The search strategy was intentionally broad to capture articles likely to report BAL cellular analysis in CHP and other ILDs, even if BAL results were only presented to characterise the study populations. No language, study design, or publication status restrictions were imposed on the initial search. Electronic database searches were supplemented with a manual review of bibliographies and searches of international conference proceedings from the European Respiratory Society (ERS) and the American Thoracic Society (ATS) from inception to August 2019.

Data from the systematic review and meta-analysis are available upon reasonable request, as made to the corresponding author. Individual patient data may be available from the corresponding authors of the cited studies.

Received: 30 Jan 2020 | Accepted after revision: 28 March 2020
Two authors (N. Adderley and C.J. Humphreys) independently screened and reviewed the articles, with discrepancies resolved by consensus and/or review by a third author (K.A. Johannson). Studies were eligible for inclusion if they met the following criteria: 1) they were original research; 2) they featured patients with a diagnosis of CHP (i.e. those in which the authors described the cohort as patients with "chronic" or "fibrotic" HP and/or if HP cohort data described the presence of radiological and/or histological fibrosis); 3) they reported BAL lymphocyte percentages; 4) they did not include paediatric patients (less than 18 years old); 5) they had full text available in English or French. In the event of multiple publications with overlapping study periods, we included only the study with the largest number of participants to prevent double counting of the patient cohorts.

Data extraction and quality assessment
Two authors (N. Adderley and C.J. Humphreys) extracted data independently and in duplicate using a standardised protocol and reporting forms. Data collected included details of study design, population characteristics, BAL lymphocyte percentage, BAL CD4:CD8 ratio, antigen determinate status (known or unknown) and specific antigen (if known), % predicted forced vital capacity (FVC), % predicted diffusing capacity of the lung for carbon monoxide (D\textsubscript{LCO}), and the definition of chronicity applied in the study. We contacted the corresponding authors of articles that reported summary statistics of BAL findings in CHP compared to other non-CHP ILDs and requested anonymized individual patient data (IPD). The variables requested for IPD were similar to those described above and were pooled as a single cohort. Assessment for risk of bias in individual studies was undertaken using the standard Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool [16].

Statistical analysis
We calculated pooled estimates using DerSimonian and Laird [17] random-effects models to assess the frequency and distribution of BAL lymphocytosis in CHP. Data were displayed graphically using forest plots. Pooled lymphocyte percentage estimates for CHP were compared to the pooled estimates for specific diagnostic categories using an unpaired t-test. When reported, healthy controls or acute/subacute HP patients were excluded from the aggregate analyses. As a sensitivity analysis, we calculated a pooled estimate from the subset of studies that defined CHP according to the presence of fibrosis on chest imaging and/or histopathology. Heterogeneity between studies was assessed using the I\textsuperscript{2} statistic to quantify the percentage of variation attributable to between-study differences [18]. We also conducted a sensitivity analysis of the pooled CHP lymphocyte percentage estimate using the inverse variance heterogeneity model. Similar methods were used to evaluate the CD4:CD8 ratio in CHP, comparing it to other non-CHP ILDs. IPD from contributing studies was pooled and treated as a single cohort to test the performance characteristics of lymphocyte thresholds in distinguishing patients with CHP from those with non-CHP ILD and specifically to differentiate them from IPF/non-IPF idiopathic interstitial pneumonia (IIP) patients. Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of BAL lymphocyte percentage were calculated at thresholds of >20%, >30%, >40% and >50%. The optimal cut-point of all pooled IPD was calculated using Youden's index [19]. Linear regression was used to identify variables associated with BAL lymphocytosis in univariate analysis, followed by multivariate analysis with pre-specified covariates including age, sex, smoking history, individual study, % predicted FVC and % predicted D\textsubscript{LCO}. Funnel plots were used to assess for publication bias.

Results
Study selection and individual patient data
The systematic review yielded 2500 unique references. After abstract and title screening, 390 articles underwent full text review, with 53 meeting the criteria for inclusion in the systematic review and 42 being included in the meta-analysis (figure 1). We contacted, or attempted to contact, the corresponding authors of 23 studies that reported BAL findings in CHP to request IPD (as well as at least one non-CHP ILD comparator population). IPD was obtained from eight unique studies.

Characteristics of included studies
Articles included in the systematic review included 23 retrospective studies and 30 prospective studies (see supplementary table S2). Studies originated from 16 countries, with 16 studies originating from Japan. Thirty studies reported data from patients with CHP only and 23 included comparator populations of non-CHP ILD. IPF was the non-CHP ILD population most frequently reported, followed by sarcoidosis. Surgical lung biopsy or high-resolution computed tomography (HRCT) findings of pulmonary fibrosis (PF) were used to define HP as chronic in 35 studies, while other methods of defining CHP were varied and are summarised in supplementary table S3. Twenty-one studies reported BAL CD4:CD8 ratio in a usable format, with a total of 315 CHP and 85 non-CHP ILD patients. IPD was obtained from eight studies, yielding a total cohort of 716 patients (188 CHP and 528 non-CHP ILD patients (229 IPF, 126
non-IPF IIP, 105 connective-tissue disease associated ILD (CTD-ILD) and 68 sarcoidosis) (table 1). Most studies were considered to be of low quality for addressing the question of how BAL lymphocytosis informs the diagnosis of CHP (supplementary table S4), with a serious risk of incorporation bias (in that BAL lymphocyte percentage was used as part of the diagnostic evaluation). Visual assessment of funnel plots demonstrated asymmetry, suggesting publication bias in the parent studies addressing BAL lymphocytosis in CHP (supplementary figure S1).

**BAL lymphocytosis in CHP versus non-CHP ILDs**

BAL lymphocyte data was extracted from a total of 42 studies with 52 unique entries (some studies reported CHP data by specific phenotypes or radiological patterns). The pooled estimate for BAL lymphocyte percentage in CHP was 42.8% (95% confidence interval (CI) 37.7–47.8, I²=95.3%) (figure 2). For IPF it was calculated (from 11 studies) as 10.0% (95% CI 6.9–13.1, I²=91.2%), while for non-IPF IIP (from five studies) it was 23.1% (95% CI –3.0 to 43.2, I²=85.2%). For CTD-ILD (from three studies) it
was calculated as 23.4% (95% CI 11.0–35.9, I²=45.7%) and for sarcoidosis (from nine studies) it was 31.2% (95% CI 17.6–44.8, I²=95.2%) (figures 3a–3d). The I² values suggest high heterogeneity for the CHP, IPF, non-IPF IIP and sarcoidosis estimates. The BAL lymphocyte percentage differed between CHP and IPF (p<0.0001), and between CHP and non-IPF IIP (p=0.0309), but not between CHP and sarcoidosis (p<0.0966). Although the number of studies was small, we identified a numerical difference between CHP and CTD-ILD (p=0.0824). The sensitivity analysis using 26 studies that defined CHP based on radiographic and/or histologic fibrosis provided a similar pooled estimate for CHP lymphocyte percentage (43.9%, 95% CI 37.4–50.4; I²=95.3%). Results were similar for the CHP pooled estimate using the inverse variance heterogeneity model (supplemental figure S2).

CD4: CD8 ratio for CHP versus other ILDs
BAL CD4:CD8 ratio data for CHP was extracted from a total of 21 studies, with a pooled estimate of 1.6 (95% CI 1.1–2.1, I²=89.3%) (figure 4). The pooled estimate for IPF was estimated (from four studies) as 1.6 (95% CI 0.6–2.5, I²=90.9%) and for sarcoidosis (from five studies) as 4.6 (95% CI 1.9–7.3, I²=87.0%). Again, the I² values suggest high heterogeneity for the CHP, IPF and sarcoidosis estimates. The CD4:CD8 ratio did not differ between CHP and IPF (p=0.9053), but did differ between CHP and sarcoidosis (p=0.0007). No CD4:CD8 data were reported for CTD-ILD.

Performance characteristics of lymphocyte thresholds
The IPD data from eight studies was pooled and analysed as a single cohort to calculate the performance characteristics of BAL lymphocytosis at four different thresholds (table 2). A comparison of studies providing IPD to those not providing IPD is presented in the supplementary material (supplementary table S5).

To identify CHP from non-CHP ILD, different threshold values yielded the following sensitivities and specificities: >20% (sensitivity=68.1%, specificity=64.8%); >30% (sensitivity=54.8%, specificity=78.9%); >40% (sensitivity=43.1%, specificity=85.5%); >50% (sensitivity=30.7%, specificity=92.4%). The PPV increased and the NPV decreased with increasing threshold values. The thresholds that maximised sensitivity and specificity were 20% and 50%, respectively. As expected, increasing specificity lowered sensitivity and vice versa. The BAL lymphocytosis value that concurrently optimised sensitivity (66.5%) and specificity (65.9%) was 21.3%. In comparison to findings from the pooled non-CHP population, BAL lymphocyte percentage differentiated patients with CHP from those with IPF/non-IPF IIP more accurately, with higher specificity and PPV (table 3). The value that optimised sensitivity (70.7%) and specificity (67.6%) was similar at 21%.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>CHP (n=188)</th>
<th>Non-CHP ILD (n=528)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age years</td>
<td>60.3±12.9</td>
<td>58.5±13.1</td>
</tr>
<tr>
<td>Male sex</td>
<td>81 [43.1]</td>
<td>284 [54.8]</td>
</tr>
<tr>
<td>Never smoker</td>
<td>66 [49.6]</td>
<td>101 [48.3]</td>
</tr>
<tr>
<td>Former smoker</td>
<td>61 [45.9]</td>
<td>79 [37.8]</td>
</tr>
<tr>
<td>Current smoker</td>
<td>6 [4.5]</td>
<td>29 [13.9]</td>
</tr>
<tr>
<td>FVC % predicted</td>
<td>68.1±19.6 [n=172]</td>
<td>73.1±20.7 [n=513]</td>
</tr>
<tr>
<td>DLCO % predicted</td>
<td>52.1±18.9 [n=158]</td>
<td>59.2±22.0 [n=435]</td>
</tr>
<tr>
<td>BAL lymphocytes %</td>
<td>35.4±24.2</td>
<td>19.8±19</td>
</tr>
<tr>
<td>BAL lymphocytes %</td>
<td>32.5 [13.6–53.5]</td>
<td>13 [6–28]</td>
</tr>
<tr>
<td>Antigen known</td>
<td>168 [89]</td>
<td>–</td>
</tr>
<tr>
<td>Antigen unknown</td>
<td>10 [5.3]</td>
<td>–</td>
</tr>
<tr>
<td>Antigen not reported</td>
<td>10 [5.3]</td>
<td>–</td>
</tr>
<tr>
<td>IPF</td>
<td>–</td>
<td>229 [43.4]</td>
</tr>
<tr>
<td>Non-IPF IIP</td>
<td>–</td>
<td>126 [23.9]</td>
</tr>
<tr>
<td>CTD-ILD</td>
<td>–</td>
<td>105 [19.9]</td>
</tr>
<tr>
<td>Sarcoidosis</td>
<td>–</td>
<td>68 [12.9]</td>
</tr>
</tbody>
</table>

Data are presented as n (%), mean±SD or median (IQR). CHP: chronic hypersensitivity pneumonitis; ILD: interstitial lung disease; FVC: forced vital capacity; DLCO: diffusing capacity of the lung for carbon monoxide; BAL: bronchoalveolar lavage; IPF: idiopathic pulmonary fibrosis; IIP: idiopathic interstitial pneumonia; CTD-ILD: connective-tissue disease associated ILD; IQR: interquartile range.
Study & Mean & 95% CI & n & Weight %
--- & --- & --- & --- & ---
Bencic 1990 & 14.10 & (6.55–21.65) & 8 & 2.0
Sterclova 2013 & 19.50 & (8.03–30.97) & 14 & 1.9
Murayama 1993 & 20.00 & (8.31–31.69) & 9 & 1.9
Ochi 2017 “concordant UIP” & 20.10 & (8.73–31.47) & 9 & 1.9
Nukui 2019 & 20.40 & (13.16–27.64) & 32 & 2.0
Nunes 2015 & 20.64 & (10.36–30.92) & 14 & 2.0
Sterclova 2006 & 23.10 & (10.14–36.06) & 7 & 1.9
Ohtani 2003 “insidious BFL” & 23.20 & (14.60–31.80) & 17 & 2.0
Unoura 2011 & 23.29 & (7.69–38.89) & 11 & 1.8
Tzilas 2019 & 24.10 & (19.00–29.20) & 35 & 2.1
Sterclova 2009 & 24.60 & (11.86–37.34) & 16 & 1.9
Vourlekis 2002 & 24.67 & (9.96–39.38) & 3 & 1.8
Ochi 2017 “discordant UIP” & 24.70 & (14.12–35.28) & 20 & 2.0
Selman 1991 “worsened” & 26.50 & (19.15–33.85) & 8 & 2.0
Inase 2006 & 26.90 & (15.36–38.44) & 3 & 1.9
Markert 2009 & 29.70 & (14.14–45.26) & 7 & 1.8
Salisbury 2018 & 29.80 & (20.34–39.26) & 16 & 2.0
Ochi 2017 “fibrotic NSIP” & 32.90 & (20.16–45.64) & 16 & 1.9
Schmidt 2002 & 34.44 & (24.15–44.73) & 5 & 2.0
Gaxiola 2011 “UIP” & 36.10 & (21.91–50.29) & 10 & 1.9
Wang 2009 & 36.60 & (21.04–52.16) & 7 & 1.8
Selman 1991 “healed” & 37.20 & (23.12–51.28) & 7 & 1.9
Remy–Jardin 1993 & 38.00 & (28.20–47.80) & 11 & 2.0
Voisin 1981 & 39.67 & (33.44–45.89) & 3 & 2.1
Kishi 2008 “fibrotic NSIP” & 40.80 & (27.91–53.69) & 10 & 1.9
Delacroix 1985 & 41.00 & (24.22–57.78) & 4 & 1.8
Caillaud 2012 & 42.60 & (35.99–49.21) & 41 & 2.0
Selman 1991 “improved” & 43.70 & (29.57–57.83) & 10 & 1.9
Bellanger 2016 & 46.00 & (31.79–60.21) & 16 & 1.9
Milanowski 1998 & 46.19 & (36.93–55.44) & 8 & 2.0
Inase 2007 & 48.90 & (27.70–70.10) & 8 & 1.6
Groot Kormelink 2011 & 49.95 & (42.37–57.54) & 22 & 2.0
Lacasse 2003 & 51.00 & (45.19–56.81) & 55 & 2.1
Gaxiola 2011 “NSIP” & 52.10 & (42.61–61.59) & 22 & 2.0
Sumi 2003 & 53.00 & (4.01–101.99) & 2 & 0.8
Koshel 2010 & 53.80 & (36.94–70.66) & 6 & 1.8
Haslam 1987 & 54.51 & (41.94–67.08) & 15 & 1.9
Pesci 1993 & 56.40 & (46.79–65.81) & 15 & 2.0
Barrera 2008 & 56.70 & (50.51–62.89) & 30 & 2.1
Garcia de Alba 2015 & 57.00 & (49.55–64.45) & 20 & 2.0
Reynolds 1977 & 62.00 & (57.07–66.93) & 10 & 2.1
Gaxiola 2011 “typical pattern” & 64.60 & (59.22–69.98) & 58 & 2.1
Leatherman 1984 & 65.00 & (57.24–72.76) & 6 & 2.0
Dai 2005 & 66.80 & (58.77–74.83) & 12 & 2.0
Ohshima 2009 & 67.00 & (62.93–71.07) & 3 & 2.1
Kishi 2008 “cellular NSIP/OP” & 67.20 & (42.80–91.60) & 6 & 1.5
Ohtani 2003 “recurrent BFL” & 69.50 & (61.30–77.70) & 15 & 2.0
Ochi 2017 “cellular NSIP” & 69.70 & (55.98–83.42) & 7 & 1.9
Tsubishima 2006 & 70.50 & (63.19–77.81) & 22 & 2.0
Ye 2009 & 77.00 & (73.08–80.92) & 16 & 2.1
Pardo 2000 & 78.90 & (70.20–87.60) & 15 & 2.0

Overall effect & 42.75 & (37.69–47.81) & 100.0 &
Prediction interval & (3.24–82.26) &

**FIGURE 2** Pooled estimate for bronchoalveolar lavage (BAL) lymphocyte percentage in chronic hypersensitivity pneumonitis (CHP). CI: confidence interval; UIP: usual interstitial pneumonia; BFL: bird fancier’s lung; NSIP: nonspecific interstitial pneumonia; OP: organising pneumonia. For details of the studies, see supplementary material.

In univariate analyses, older age, male sex and ever having smoked were associated with lower lymphocyte percentage in patients with CHP (table 4). In pre-specified multivariate analysis, age ($\beta=-0.32$, 95% CI $-0.58$ to $-0.06$; $p=0.016$) and ever having smoked ($\beta=-11.3$, 95% CI $-19.9$ to $-2.6$; $p=0.011$) were
FIGURE 3 Pooled estimates for bronchoalveolar lavage (BAL) lymphocyte percentage in (a) idiopathic pulmonary fibrosis (IPF), (b) non-IPF idiopathic interstitial pneumonia (IIP), (c) connective-tissue disease associated ILD (CTD-ILD) and (d) sarcoidosis. CI: confidence interval. For details of the studies, see supplementary material.
associated with lower lymphocyte percentage in patients with CHP. In multivariate analysis, we did not identify an association between lung function measures and BAL lymphocyte percentage, although these may not accurately reflect fibrosis extent or disease severity in this cohort. Of 169 CHP patients with an implicated antigen exposure, 126 (75%) were reportedly bird fanciers, while the remaining 43 patients (25%) had diverse other exposures. We did not identify an association between antigen type and BAL lymphocyte percentage, comparing bird fanciers to other exposures.

Discussion

BALF lymphocyte percentage is higher in patients with CHP compared to other non-CHP ILDs, most notably IPF and non-IPF IIP. However, there is high heterogeneity across studies and, given that almost all studies used BALF cellular analysis findings as part of the CHP diagnostic evaluation, there exists a high

<table>
<thead>
<tr>
<th>Study</th>
<th>Mean</th>
<th>95% CI</th>
<th>n</th>
<th>Weight %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sumi 2003</td>
<td>0.23</td>
<td>[0.07–0.39]</td>
<td>2</td>
<td>7.7</td>
</tr>
<tr>
<td>Bencic 1990</td>
<td>0.37</td>
<td>[0.13–0.61]</td>
<td>8</td>
<td>7.5</td>
</tr>
<tr>
<td>Schmidt 2002</td>
<td>0.51</td>
<td>[0.08–0.94]</td>
<td>5</td>
<td>6.9</td>
</tr>
<tr>
<td>Tsushima 2006</td>
<td>0.70</td>
<td>[0.57–0.83]</td>
<td>22</td>
<td>7.8</td>
</tr>
<tr>
<td>Leatherman 1984</td>
<td>0.80</td>
<td>[0.64–0.96]</td>
<td>6</td>
<td>7.7</td>
</tr>
<tr>
<td>Ohtani 2003 “recurrent BFL”</td>
<td>0.85</td>
<td>[0.46–1.24]</td>
<td>15</td>
<td>7.0</td>
</tr>
<tr>
<td>Koschel 2010</td>
<td></td>
<td>[-0.06 to 2.06]</td>
<td>6</td>
<td>4.1</td>
</tr>
<tr>
<td>Sterclova 2006</td>
<td>1.40</td>
<td>[0.51–2.29]</td>
<td>7</td>
<td>4.8</td>
</tr>
<tr>
<td>Vourlekis 2002</td>
<td>1.40</td>
<td>[1.20–1.60]</td>
<td>2</td>
<td>7.7</td>
</tr>
<tr>
<td>Tzilas 2019</td>
<td>1.42</td>
<td>[1.04–1.80]</td>
<td>13</td>
<td>7.1</td>
</tr>
<tr>
<td>Nunes 2014</td>
<td>2.06</td>
<td>[0.99–3.13]</td>
<td>4</td>
<td>4.1</td>
</tr>
<tr>
<td>Sterclova 2009</td>
<td>2.10</td>
<td>[0.92–3.28]</td>
<td>16</td>
<td>3.7</td>
</tr>
<tr>
<td>Unoura 2011</td>
<td>2.40</td>
<td>[1.04–3.76]</td>
<td>11</td>
<td>3.2</td>
</tr>
<tr>
<td>Murayama 1993</td>
<td>2.60</td>
<td>[0.90–4.30]</td>
<td>9</td>
<td>2.4</td>
</tr>
<tr>
<td>Ochi 2017 “fibrotic NSIP”</td>
<td>2.60</td>
<td>[-2.30 to 7.50]</td>
<td>16</td>
<td>0.4</td>
</tr>
<tr>
<td>Kishi 2008 “fibrotic NSIP”</td>
<td>2.74</td>
<td>[0.82–4.64]</td>
<td>10</td>
<td>2.0</td>
</tr>
<tr>
<td>Kishi 2008 “cellular NSIP/OP”</td>
<td>2.77</td>
<td>[-0.91 to 6.45]</td>
<td>6</td>
<td>0.7</td>
</tr>
<tr>
<td>Dai 2005</td>
<td>3.30</td>
<td>[0.56–6.04]</td>
<td>12</td>
<td>1.1</td>
</tr>
<tr>
<td>Ochi 2017 “cellular NSIP”</td>
<td>3.30</td>
<td>[-5.32 to 11.92]</td>
<td>7</td>
<td>0.1</td>
</tr>
<tr>
<td>Barrera 2008</td>
<td>4.38</td>
<td>[2.91–5.85]</td>
<td>30</td>
<td>2.9</td>
</tr>
<tr>
<td>Ochi 2017 “concordant UIP”</td>
<td>4.40</td>
<td>[-1.48 to 10.28]</td>
<td>9</td>
<td>0.3</td>
</tr>
<tr>
<td>Nukui 2019</td>
<td>4.70</td>
<td>[2.94–6.46]</td>
<td>22</td>
<td>2.3</td>
</tr>
<tr>
<td>Ohtani 2003 “insidious BFL”</td>
<td>5.05</td>
<td>[2.35–7.75]</td>
<td>17</td>
<td>1.1</td>
</tr>
<tr>
<td>Ochi 2017 “discordant UIP”</td>
<td>5.20</td>
<td>[-6.36 to 16.76]</td>
<td>20</td>
<td>0.1</td>
</tr>
<tr>
<td>Kishi 2008 “UIP”</td>
<td>5.86</td>
<td>[2.97–8.75]</td>
<td>16</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Overall effect                  1.58  [1.06–2.11] 100.0

Prediction interval

Heterogeneity: I²=89%

| TABLE 2 Performance characteristics at lymphocyte percentage thresholds to identify chronic hypersensitivity pneumonitis [CHP] from non-CHP interstitial lung disease [ILD] |
|---------------------------------|--------|----------------|--------|-----------|
| Threshold % | Sensitivity % | Specificity % | PPV %  | NPV %     |
| 20         | 68.1 [60.9–74.7] | 64.8 [60.5–68.9] | 40.8 [37.2–44.5] | 85.1 [82.1–87.6] |
| 30         | 54.8 [47.4–62.0] | 78.9 [75.3–82.2] | 46.6 [41.6–51.7] | 82.9 [80.4–85.1] |
| 40         | 43.1 [35.9–50.5] | 85.5 [82.2–88.4] | 51.6 [45.0–58.14] | 80.7 [78.7–82.7] |
| 50         | 30.7 [24.3–37.8] | 92.4 [89.8–94.5] | 59.6 [50.6–68.0] | 78.4 [76.8–80.1] |

Data are presented as n (95% CI). PPV: positive predictive value; NPV: negative predictive value; CI: confidence interval.
predicted FVC nor % predicted lymphocyte percentage, with smoking status known to influence BAL cellular analyses [5]. Neither %
acute HP [11, 25, 26]. In our IPD analysis, older age and a history of smoking were associated with lower between them. Data suggest that the extent of PF, the type of antigen and the time since last exposure will
characteristics of the studies, we could not perform robust subset analyses to explain the differences thresholds (providing higher specificity) may be more clinically useful. Older age and ever having smoked
were associated with lower BAL lymphocyte percentage in CHP. Unfortunately, the data do not inform the impact of other clinical variables on BAL lymphocyte percentage in a patient with suspected CHP.

CHP can be challenging to diagnose, largely due to the historical lack of widely accepted consensus criteria. The role of BAL lymphocyte percentage in establishing a diagnosis of CHP has remained controversial, with its use largely carried over from clinical experience with acute (non-fibrotic) HP [4, 20, 21]. The most recent international IPF guidelines [22] conditionally recommended BAL in patients with suspected IPF and a non-diagnostic HRCT pattern, although a meta-analysis of eight studies reported therein found no difference in BAL lymphocyte percentage between IPF and CHP. Notably, the HP BAL lymphocytosis data used in this analysis was drawn from only two studies, both excluded from our analysis due to patients not having “chronic” or “fibrotic” HP, or the full text being unavailable in English or French [23, 24]. Clinicians and patients must consider the risk–benefit ratio of diagnostic tests and the anticipated yield of clinically relevant information. Our findings highlight that no studies have robustly tested the additive discriminative value of BAL lymphocytosis in differentiating CHP from other forms of fibrotic ILD. The identification of a threshold that provides the best sensitivity and specificity suggests that such a value could be tested prospectively to determine the performance of BAL cellular analysis as a diagnostic test in CHP. However, given the limitations of the data, this number should not as of yet be considered a diagnostic test, either to rule in or rule out CHP.

The pooled estimates for BAL lymphocyte percentage show high heterogeneity and, based on the characteristics of the studies, we could not perform robust subset analyses to explain the differences between them. Data suggest that the extent of PF, the type of antigen and the time since last exposure will influence the degree of alveolar inflammation in HP, with lower lymphocyte percentage in CHP relative to acute HP [11, 25, 26]. In our IPD analysis, older age and a history of smoking were associated with lower lymphocyte percentage, with smoking status known to influence BAL cellular analyses [5]. Neither % predicted FVC nor % predicted $D_{\text{LCO}}$ were associated with degree of lymphocytosis, although these may not be accurate surrogates of fibrosis. The lack of granular data on antigen type and time since exposure limited our ability to evaluate associations with lymphocytosis and this is an area in need of further study. The CD4:CD8 ratio was found to be lower in CHP than in sarcoidosis, albeit with high between-study

<table>
<thead>
<tr>
<th>Variable</th>
<th>Univariate analysis</th>
<th>Multivariate analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta$-coefficient (95% CI)</td>
<td>p-value</td>
</tr>
<tr>
<td>Age</td>
<td>$-0.42 [-0.68 to -0.15]$</td>
<td>0.002</td>
</tr>
<tr>
<td>Sex</td>
<td>$-12.9 [-19.7 to -6.07]$</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Smoking</td>
<td>$-16.0 [-22.9 to -9.1]$</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>FVC %</td>
<td>$-0.3 [-0.48 to -0.12]$</td>
<td>0.001</td>
</tr>
<tr>
<td>$D_{\text{LCO}}$ %</td>
<td>$-0.1 [-0.3 to 0.1]$</td>
<td>0.306</td>
</tr>
</tbody>
</table>

CI: confidence interval; FVC: forced vital capacity; $D_{\text{LCO}}$: diffusing capacity of the lung for carbon monoxide.
variability. The CD4:CD8 ratio is known to be influenced by exposure, smoking and disease severity, and this test has largely fallen out of favour in clinical practice [5, 27].

This study has important limitations. Despite the breadth of our literature search, relevant references may have been missed. We specifically focused on chronic/fibrotic HP and therefore these findings cannot be extrapolated to acute/non-fibrotic forms of HP. Given the heterogeneity in diagnostic criteria for CHP across studies, we cannot determine the validity of the CHP diagnoses. We tried to address this through the sensitivity analysis that demonstrated consistent findings. Data from the parent studies did not permit assessment of BAL technique and quality, or the potential impact of treatment on BAL cellular analysis. Furthermore, the data did not allow stratification of patients based on degree of fibrosis or presence of other HRCT morphologic features (e.g. air-trapping, centrilobular nodules and ground glass opacities) and this may impact the test performance characteristics. In addition, our data did not allow for strong characterisation of antigen exposure, a variable that likely influences the degree of alveolitis. Most importantly, the validity of our findings is limited by the quality of the parent studies, the majority of which are subject to incorporation bias. Despite these limitations, our study has several important strengths. We identified studies with an intentionally broad search strategy and created a large cohort using individual patient data. To the best of our knowledge, the current study provides the most comprehensive assessment of this clinically meaningful question in this patient population to date.

Conclusions
BAL lymphocyte percentage is higher in CHP compared to non-CHP forms of ILD, with older age and ever smoking associated with lower lymphocyte percentage. Higher thresholds of lymphocyte percentage provide greater specificity at a cost of sensitivity. Further work is needed to inform the role of BAL in the absence of incorporation bias, by testing the discriminative performance of lymphocyte percentage in established diagnostic prediction models. Finally, a deeper understanding of the relationships between antigen exposure, host factors and alveolar lymphocytosis will guide the use of BAL in the diagnostic evaluation of CHP.

Author contributions: N. Adderley and K.A. Johannson conceived the study; all authors contributed to the study design and to protocol development; H. Barnes and K.A. Johannson conducted the statistical analyses; N. Adderley and K.A. Johannson drafted the manuscript; all authors contributed to, critically appraised and approved the final version of the manuscript.

Conflict of interest: N. Adderley has nothing to disclose. C.J. Humphreys has nothing to disclose. H. Barnes has nothing to disclose. B. Ley has nothing to disclose. Z.A. Premji has nothing to disclose. K.A. Johannson reports personal fees for advisory board work, consultancy, lectures and travel to meetings from Boehringer Ingelheim, personal fees for advisory board work and lectures from Hoffman La Roche Ltd, personal fees for advisory board work and consultancy from Theravance and Blade Therapeutics, grants from the Chest Foundation, the University of Calgary School of Medicine, the Pulmonary Fibrosis Society of Calgary and UCB Biopharma SPRL, outside the submitted work.

References